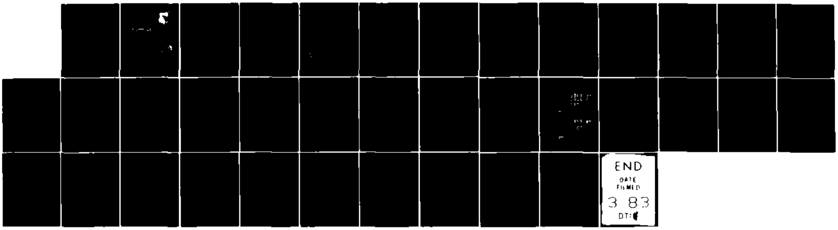
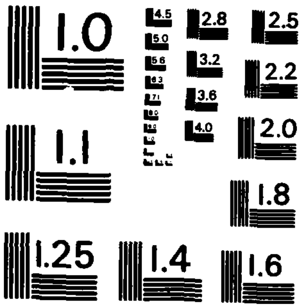
AD A124 307 VISUAL JUDGMENTS OF OPTICAL DISTORTIONS IN AIRCRAFT 1/1
WINDSCREENS(U) AIR FORCE AEROSPACE MEDICAL RESEARCH LAB
WRIGHT-PATTERSON AFB OH H C SELF OCT 82

UNCLASSIFIED AFAMRL-TR-81-24 F/G 1/3 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFAMRL-TR-81-24

ADA124307

VISUAL JUDGMENTS OF OPTICAL DISTORTIONS IN AIRCRAFT WINDSCREENS

HERSCHEL C. SELF, Ph. D.

AIR FORCE AEROSPACE MEDICAL RESEARCH LABORATORY

OCTOBER 1982

AIR FORCE AEROSPACE MEDICAL RESEARCH LABORATORY
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from Air Force Aerospace Medical Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
Cameron Station
Alexandria, Virginia 22314

TECHNICAL REVIEW AND APPROVAL

AFAMRL-TR-81-24

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Chief
Human Engineering Division
Air Force Aerospace Medical Research Laboratory

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFAMRL-TR-81-24	2. GOVT ACCESSION NO.: A124307 / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): VISUAL JUDGMENTS OF OPTICAL DISTORTIONS IN AIRCRAFT WINDSCREENS	5. TYPE OF REPORT & PERIOD COVERED: Technical Report 6 Oct 76 - 18 Jan 81
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Hershel C. Self, Ph. D.	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Air Force Aerospace Medical Research Laboratory Aerospace Medical Division, Air Force Systems Command, Wright-Patterson AFB OH 45433	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62202F, 7184-18-03
11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE: October 1982
	13. NUMBER OF PAGES: 36
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Distortion, Windscreens, Windshields, Tests, Inspection, Evaluation, Operators, Vision, Ergonomics, Human Engineering, Quality Inspection, Acceptance Testing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Observer ratings of optical distortion in eleven F-111 Aircraft windscreens were examined using six factory production line visual quality inspectors, six Air Force pilots, all with years of flying experience, and two observers familiar with aircraft windscreen problems. Observers looked through the windscreens at large gridboards having thin white lines on a black background. Each windscreen was rated for effect of distortion on flying performance (yes-no), acceptability (yes-no), and for position on a 0-5 distortion scale for eight optical distortion variables: line splitting, line bending, line banding, shimmer,

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

magnification, other distortions, and overall distortion.

High correlations were found between types of distortion. Ratings on either banding or line bending, could be used to efficiently predict overall optical distortion. Pilots and visual quality inspectors were quite close in judgements of overall optical distortion and on specific types of distortion. Pilots rated distortions very slightly worse (higher), but were appreciably more likely to rate a windscreen as influencing pilot performance. Neither lens factor nor displacement grade, alone, were significantly related to acceptability or performance effects judgments.

SUMMARY OF RESULTS

1. For the thick multilayer plastic F-111 windscreens used in the present study there was an appreciable positive correlation between lens factor (LF) and displacement grade (DG). When, for a windscreen, either one is large, the other also tends to be large, and similarly for intermediate and low values.

2. There was no significant amount of relationship between either LF alone or DG alone and either the percentage of observes rating the windscreens as acceptable (% A) or the percentage who said that enough distortion was present to influence flying performance (% I).

3. Neither K. E. Thompson's windscreen, acceptability graph based on a combination of LF and DG nor the present author's "R" criterion of $\sqrt{(LF)^2 + (DG/1{,}000 + 1)^2}$ was effective in sorting windscreens according to either Acceptability (% A) or Estimated Influence (% I).

4. As expected, the correlation between % I and % A was high, negative and statistically significant. When either is high for a windscreen, the other tends to be low.

5. There were high positive statistically significant correlations between estimated overall distortion and subjective estimates of shimmer, grid line banding, grid line blur and grid line bending, but not with grid line splitting. Splitting did not correlate significantly with any other optical distortion measure.

6. The types of optical distortion that have a large affect upon acceptability (% A), estimated influence (% I) and estimated overall optical distortion were shimmer, banding, blurring and bending.

7. There were high (r = .61-.93) positive statistically significant correlations between shimmer, banding and blur estimates.

8. Grid line banding ratings correlate +.961 and grid line bending ratings, +.932 with estimated overall optical distortion. Thus, ratings on either banding or bending could be used to predict overall optical distortion ratings.

9. The multiple correlation coefficient, R, between overall optical distortion ratings and predicted overall optical distortion ratings based on the estimates of shimmer, banding, blurring and bending is +.9763.

10. An equation for predicting estimated overall distortion from estimated shimmer, banding, blurring and bending did an excellent job of prediction, as such a large R would indicate. The percentage error in predicting was less than 10% for 9 out of 11 windscreens: one windscreen prediction was in error by 10.6% and another erred by 17.4%.

11. Out of the 14 subjects (or observers) in the present study, 6 were Pittsburgh Plate Glass Co. visual quality inspectors. When distortion ratings of the two groups were compared it was found that: On six of the seven types of optical distortion, agreement between pilots and inspectors were quite close. Even on the one exception (grid line splitting), both groups rated closer to "Zero" distortion, the lowest category, than to "Little" distortion. For all practical purposes, agreement was close on all types of optical distortion. The r between the two groups on estimates of all seven distortions was +.9557.

12. Although differences between pilots and inspectors are negligible, the inspectors tended to rate optical distortion slightly lower than did pilots on all types of distortion.

13. In terms of use of rating categories, the inspectors tended to use the "No Distortion" category significantly more often than the pilots did, with a consequent tendency to use the "Little Distortion" category less often. However, agreement between pilots and inspectors was close on the "Moderate" and "Large" distortion categories, and the magnitude of the difference in the "Extreme" distortion category was small.

14. In terms of particular windscreens, the average ratings for the two groups on the two windscreens for which pilot and inspector ratings differed the most were actually only a fraction of an optical distortion category apart. The inspectors were not more lenient (less critical) in their ratings than pilots were for the higher distortion windscreens for which rating differences could have a practical significance.

15. Pilots were more critical than inspectors in judging the influence on flying (% I) of optical distortion and in rating windscreens on acceptability (% A). The differences were statistically significant. Pilots and visual quality inspectors, while agreeing fairly closely on amounts of several types of distortion and on the overall optical distortion of specific windscreens, were not very close in judging (or rating) acceptability (% A) and influence upon flying of distortions (% I).

PREFACE

This study was part of a joint effort by the Crew Station Integration and Technology Development Branches of the Human Engineering Division, Air Force Aerospace Medical Research Laboratory (AFAMRL) at Wright-Patterson Air Force Base, Ohio. It was done under Project 7184, Man-Machine Integration Technology, Task 18, Visual Effects of Windscreens on Pilot Performance, and Work Unit 03, Visual Perception Through Windscreens. The work was undertaken in support of the Improved Windscreen Protection Development Program, Projects 2202 and 1926 of the Air Force Wright Aeronautical Laboratories, Flight Dynamics Laboratory (AFWAL/FD) Wright-Patterson Air Force Base, Ohio.

The report analyzes data collected at the Pittsburgh Plate Glass Company (PPG), Creighton, Pennsylvania. The courtesy, cooperation and support rendered by the management and personnel of the company are acknowledged. The author also is thankful for the services of the USAF pilots and the PPG visual quality inspectors who made the study possible. The assistance of Lt Col Joseph A. Birt in arranging for the test subjects and securing the cooperation of the manufacturer is appreciated.

This report covers one of the two studies conducted at the PPG company facility using the same pilots and inspectors as subjects.

TABLE OF CONTENTS

Section *Page*

I INTRODUCTION 6
PROBLEM AND OBJECTIVE 6
WINDSCREEN DISTORTION 6

II OBSERVERS, INSTRUCTIONS AND TESTING 9
PARTICIPATING OBSERVERS 9
INSTRUCTIONS TO OBSERVERS 9
THE DISTORTION RATING SHEETS 9

III RESULTS 11
LENS FACTOR AND DISPLACEMENT GRADE 11
ESTIMATED INFLUENCE AND ACCEPTABILITY 15
ESTIMATED INFLUENCE AND DISTORTION ESTIMATES 15
ACCEPTANCE AND DISTORTION ESTIMATES 15
OVERALL DISTORTION AND DISTORTION ESTIMATES 15
SOME RELATIONSHIPS BETWEEN TYPES OF DISTORTION 19
DISTORTION RATINGS OF PILOTS AND QUALITY INSPECTORS 21
CATEGORIES OF DISTORTION USED BY PILOTS AND INSPECTORS 25
INDIVIDUAL WINDSCREEN RATINGS BY PILOTS AND VQI 25
PILOTS AND INSPECTORS ON INFLUENCE AND ACCEPTABILITY 30

IV DISCUSSION OF RESULTS 32

LIST OF ILLUSTRATIONS

Figure *Page*

1. THE SCORE SHEET FORMAT 10
2. PLOT OF LF AND DG FOR THE ELEVEN F-111 WINDSCREENS ON A K.E. THOMPSON WINDSCREEN ACCEPTABILTY GRAPH 13
3. % I PLOTTED AGAINST DISPLACEMENT GRADE 14
4. % I PLOTTED AGAINST LENS FACTOR 14
5. AVERAGE DISTORTION RATINGS OF 6 PILOTS AND 6 VISUAL QUALITY INSPECTORS ON ELEVEN F-111 WINDSCREENS 23
6. RESPONSE FREQUENCIES IN VARIOUS INTERVALS OF THE DISTORTION SCALE FOR PILOTS AND VISUAL QUALITY INSPECTORS 26
7. AVERAGE DISTORTION RATINGS OF ELEVEN F-111 WINDSCREENS FOR PILOTS AND VISUAL QUALITY INSPECTORS 28
8. DISTORTION RATINGS FOR PILOTS AND INSPECTORS PLOTTED AGAINST AMOUNT OF DISTORTION ESTIMATED AS THE PILOT-INSPECTOR AVERAGE 29

LIST OF TABLES

Table *Page*

1. LENS FACTOR, DISPLACEMENT GRADE, THEIR COMPOSITE, AND THE ACCEPTABILITY TO PILOTS OF 65 F-111 WINDSCREENS (K.E. THOMPSON DATA) 8
2. DISTORTION DATA FOR 11 F-111 WINDSCREENS 12
3. PRODUCT MOMENT CORRELATIONS BETWEEN MEASURES FOR ALL 11 F-111 WINDSCREENS 16
4. PRODUCT MOMENT CORRELATIONS BETWEEN MEASURES FOR ALL WINDSCREENS EXCEPT WINDSCREEN "M" 17
5. 100 X (COEFFICIENT OF DETERMINATION) FOR STATISTICALLY SIGNIFICANT CORRELATION COEFFICIENTS:
   (a) ALL 11 F-111 WINDSCREENS
   (b) ALL WINDSCREENS EXCEPT WINDSCREEN "M" 18
6. CORRELATIONS BETWEEN DISTORTION TYPES FOR ALL 11 F-111 WINDSCREENS 19
7. OBTAINED AND PREDICTED OVERALL DISTORTION RATINGS 20
8. STANDARD PARTIAL REGRESSION COEFFICIENTS FOR PREDICTING OVERALL WINDSCREEN DISTORTION 20
9. AVERAGE DISTORTION RATINGS FOR PILOTS AND VISUAL QUALITY INSPECTORS ... 22
10. FREQUENCIES OF USE OF RATING CATEGORIES: (A) BY TYPE OF DISTORTION 24
11. NUMBER OF RESPONSES BY DISTORTION CATEGORY FOR PILOTS AND FOR VISUAL QUALITY INSPECTORS 25
12. FREQUENCIES OF USE OF RATING CATEGORIES (B) BY INDIVIDUAL WINDSCREEN ... 27
13. PERFORMANCE INFLUENCE AND ACCEPTABILITY 31
14. PERFORMANCE INFLUENCE 31
15. ACCEPTABILTY 31

SECTION I

INTRODUCTION

PROBLEM AND OBJECTIVE

It is difficult to fabricate aircraft windscreens that do not contain some optical distortion. This is particularly true when windscreens are thick, of multilayer construction, highly curved, and installed so that the user looks through them at an angle departing by over 30° from normal incidence. All of these favorable-to-distortion conditions were present in some of the early F-111 aircraft windscreens. A compromise must be made between optical perfection, which is materially and economically unfeasible, and severe optical distortion that is intolerable to pilots. Some distortion is acceptable. The Air Force needs data to establish visual distortion tolerances for use in purchasing windscreens. What amount of distortion will be visually acceptable and what amount can be present without causing more than a negligible effect on piloting an aircraft?

The present study was conducted to obtain some information about human judgments of optical distortion in windscreens. The data were collected at a fabrication plant owned by the Pittsburgh Plate Glass Co. using 11 F-111 windscreens. The current (1981) F-111 windscreens are not made like the ones in this study. Part of them were right-side windscreen panels and part were left-side panels. The two are identicial except that one is "right-handed" and the other "left-handed."

Since only 11 windscreens were available for distortion testing, and the testing used stationary gridboards rather than moving or dynamic scenes, the judgments of the observers will depart somewhat from ratings made by observers of a real moving scene. Possibly more of the windscreens would not be acceptable in flight tests, especially if observation times were extended, than would be unacceptable in static tests. Due to the small sample size and the static testing, the present study is concerned largely with the interrelationships of distortion measures and comparisons of the distortion ratings of pilots and visual quality inspectors.

WINDSCREEN DISTORTION

Optical distortion in aircraft windscreens can, from the viewpoint of an observer, take many different forms or modes of appearance. For example, the lines of a grid board, when observed through an aircraft windscreen, may appear to split, to bend, to blur, to shimmer when the observer's head is moved, to show banding effects, to exhibit "bulls eyes" due to localized magnification, etc. These effects are due to variations in the thickness of materials and in the surface configurations. The windscreen acts as though it contains surface "waves", spherical, cylindrical and aspheric lenses of low refractive power, prisms of low power, etc. The distortions are refractive or light bending phenomena.

Clearly, optical distortion, as found in aircraft windscreens, is a complex of variables rather than a single variable. Unfortunately, the technology of windscreen optical distortion measurement is not well developed. Much work is being done in various laboratories, using a wide variety of approaches, to advance the state of the art. At the present time there are several objective measures of optical distortion in aircraft windscreens. Two that were used on F-111 windscreens are lens factor (LF) and displacement grade (DG). LF is a measure of lensing or magnification. It is the cube of the size increase in grid squares per inch on a photograph of a grid pattern ("grid board") taken through the windscreen with the camera located at the pilot's design eye position. One thousand times the sum of the greatest horizontal and vertical displacements of grid lines from the same photograph in the windscreen area with the largest lens factor is termed displacement grade.

We gave Mr. K. E. Thompson of General Dynamics the problem of making optical measurements of windscreens that would allow his company to detect windscreens having optical distortions that would

* Thompson K. E., *Optical Evaluation of the F-111 Windscreen*, FZM-12-10952, Feb. 12, 1970. This is an internal publication of the Fort Worth Division of General Dynamics Corporation.

result in rejection of the windscreens by aircraft pilots. In 1969* he showed that LF and DG, in combination, could be used to divide windscreens into three groups: acceptable to pilots, questionable, and not acceptable to pilots. Very few, only eight, of the 65 F-111 windscreens that he used fell into the questionable or gray area of his graph of LF versus DG. The questionable ones required visual examination to determine acceptability to aircraft pilots. Thompson's thin glass F-111 windscreens differed both mechanically and optically from the 11 thick multilayer plastic ones of the present study. Measurements on Thompson's acceptability graph yielded LF and DG values which are listed in table 1. The table also gives the calculated values of R, an LF-DG composite derived by the author of the present report to do the same task of acceptability-to-pilots discrimination as Thompson's acceptability graph. Numerically $R = \sqrt{(LF)^2 + (DG/1{,}000 + 1)^2}$. Note from the table that the values of R, listed in order of increasing numerical values may be divided into three areas. They are separated in the table. They are: an acceptable (+) range, R = 0-1.682, where all windscreens are acceptable; a questionable or gray area, R = 1.685-1.725, where some are acceptable and some are not acceptable (—); and an unacceptable area or region, R > 1.726, where all windscreens are unacceptable. Eight of Thompson's 65 windscreens are in the gray area in the table. One of the tasks of the present report is to examine the suitability of R for use with the 11 F-111 windscreens examined at the Creighton PPG plant.

TABLE 1

LENS FACTOR, DISPLACEMENT GRADE, THEIR COMPOSITE, AND THE ACCEPTABILITY TO PILOTS OF 65 F-111 WINDSCREENS (K. E. THOMPSON DATA*)

NO	LF	DG	R	A	NO	LF	DG	R	A	NO	LF	DG	R	A
1	1.041	71.2	1.493	+	24	1.197	90.0	1.619	+	43	1.335	128.2	1.748	-
2	1.047	80.0	1.504	+	23	1.150	150.5	1.627	+	46	1.329	140.8	1.751	-
3	1.047	90.5	1.512	+	25	1.197	110.0	1.632	+	47	1.331	150.5	1.759	-
4	1.048	99.5	1.519	+	26	1.208	110.0	1.641	+	48	1.329	160.0	1.764	-
6	1.038	109.8	1.520	+	27	1.200	120.5	1.642	+	49	1.331	180.0	1.779	-
7	1.047	109.8	1.526	+	28	1.209	120.0	1.648	+	50	1.341	190.0	1.793	-
5	1.058	99.5	1.527	+	29	1.208	129.5	1.654	+	51	1.347	192.5	1.799	-
8	1.050	120.0	1.535	+	30	1.208	140.0	1.661	+	52	1.400	140.0	1.805	-
9	1.056	118.0	1.538	+	31	1.208	170.5	1.682	+	53	1.400	152.9	1.818	-
10	1.053	124.2	1.540	+	33	1.269	110.8	1.685	-	58	1.329	251.5	1.826	-
13	1.100	90.8	1.549	+	34	1.277	120.0	1.699	+	54	1.408	170.8	1.831	-
11	1.053	140.0	1.552	+	35	1.269	130.0	1.699	+	55	1.400	190.0	1.837	-
14	1.093	106.0	1.555	+	32	1.197	211.0	1.703	-	57	1.397	210.8	1.848	-
12	1.058	142.0	1.557	+	36	1.259	149.5	1.705	+	59	1.472	150.0	1.868	-
16	1.100	109.5	1.562	+	37	1.269	149.8	1.712	-	60	1.471	170.0	1.880	-
15	1.103	108.0	1.563	+	38	1.269	160.0	1.719	-	56	1.460	206.8	1.894	-
17	1.100	120.0	1.570	+	39	1.277	160.0	1.725	+	62	1.500	180.5	1.909	-
18	1.150	100.0	1.591	+	40	1.330	100.0	1.726	-	61	1.480	211.2	1.912	-
19	1.150	110.8	1.599	+	41	1.321	120.0	1.732	-	63	1.550	221.5	1.973	-
20	1.150	120.0	1.605	+	44	1.324	132.5	1.742	-	64	1.620	294.8	2.074	-
21	1.153	120.0	1.607	+	42	1.338	117.8	1.743	-	65	1.819	232.2	2.197	
22	1.150	130.0	1.612	+	45	1.329	132.2	1.746	-					

*Taken from measurements made on Thompson's graph.

Key:

NO = Windscreen Number on Graph; LF = Lens Factor:

DG = Displacement Grade; A = Acceptability to Pilots;

R = LF and DG Composite = Vector Length or Distance = $\sqrt{(LF)^2 + (DG/1,000 + 1)^2}$

SECTION II

OBSERVERS, INSTRUCTIONS AND TESTING

PARTICIPATING OBSERVERS

All of the people who served as observers and raters in this study had already spent several hours in making judgments about the various defects present in the 11 windscreens that were to be examined and rated. They had all been instructed not to discuss their ratings or observations with other observers. Six of the observers (or subjects) had had considerable experience with judging the visual quality of aircraft windscreens. This group will be referred to as visual quality inspectors (VQI). It included five Pittsburgh Plate Glass visual quality inspectors and one Air Force Major with several years of experience in the laboratory and in the field with windscreen inspection. Six of the observers were Air Force pilots all of whom had several years of flying experience. The two remaining observers were familiar with aircraft windscreen problems, but were neither pilots nor visual quality inspectors.

INSTRUCTIONS TO OBSERVERS

Observers were told that they would be asked to make judgments about 11 F-111 windscreen panels, some left or pilot's panels, and some right or copilot's panels. They were to stand on a bench and look through each panel at a well-illuminated grid board pattern. The grid board consisted of tightly stretched white strings just in front of a black panel, giving white lines on a black background. The string formed 1-inch squares. Observers were permitted to move their heads in any way that they desired while making ratings. They were told that there was no time limit, and that they were not to discuss their observations or ratings with other participants. Each windscreen was rated and the rating sheet was handed in before the next windscreen was seen. The identification of individual windscreen panels was concealed by code letters so that raters would not remember previous ratings or judgments made on specific windscreens.

THE DISTORTION RATING SHEETS

As each observer stepped up to a windscreen to rate it on distortion he was handed a score sheet containing questions and blank spaces for writing down or checking his responses. The score sheets for all 11 windscreens were identical. Figure 1 is a sample blank score sheet. Note that the first question asked the observer if there was enough distortion in the windscreen to affect flying performance. The second asked if the windscreen was acceptable. Question 3 had the observer rate the windscreen on eight distortion variables: line splitting, line bending, line blurring, banding, shimmer, magnification (bulls eyes), other, and overall distortion. There were five checklist categories of amount of distortion observed: none (0), little (1), moderate (2), large (3), and extreme (4). The observer checked the space indicating the appropriate amount. A space was also provided for indicating the location of the distortion on a standard diagram of the windscreen.

The last question asked the observer, if he rejected the windscreen, to indicate, relevant to the standard diagram, the location of the area that had such bad distortion as to cause windscreen rejection.

Part Identification ____________________ Name ______________________

1. Would you say that there is enough distortion in this part to affect flying performance?

   Yes _____ No _____

2. In your opinion, is this part acceptable?

   Yes _____ No _____

3. The following optical defects are components of distortion. Please rate the part on these defects by checking the scale below (0 thru 4). Also, if a moderate amount or more is checked (category 2 or higher), indicate in the last column the area or areas (from the diagram provided) where the defect was observed.

	Amount Observed					
	None	Little	Moderate	Large	Extreme	Diagram Locations
A. Line Split						
B. Line Bending						
C. Line Blurring						
D. Banding						
E. Shimmer						
F. Magnification (Bulls Eyes)						
G. Other						
H. Overall Distortion						

4. If you rejected the part in question "2" above, what defect (A), in what area (where), on the diagram, caused your rejection?

Figure 1 The score sheet format.

SECTION III

RESULTS

LENS FACTOR AND DISPLACEMENT GRADE

Table 2, under the column heading "Objective Measures," lists the lens factor (LF) and displacement grade (DG) of the 11 multilayer plastic F-111 windscreens. Figure 2 is a plot of LF against DG. The plot includes Thompson's area of probable acceptance, area requiring subjective evaluation, and area of probable rejection. Note that windscreen "M" is atypical: it is the only one of the 11 with a high displacement grade combined with a low lens factor. The figure gives the least-squares best fit linear equation relating LF and DG for the 11 windscreens: DG = 131 LF — 28.6. The data point "M" was not used in the derivation.

The Pearson product-moment correlation coefficient, r, between LF and DG was +.6311 for all 11 windscreens and +.8752 when atypical windscreen "M" was excluded. Both of these r values are statistically significant at the .05 level of significance. There is an appreciable tendency for either value to be high when the other is high, and similarly for medium and for low values.

The fairly high positive LF-DG correlation is not surprising. The author had previously measured, from a graph in Thompson's report, the LF and DG values of the 65 glass F-111 windscreens used by K. E. Thompson in 1969 at General Dynamics and calculated an r of +.7554 between LF and DG. This large positive value is statistically significant at the .01 level of statistical significance. Thus, LF and DG vary together in both glass and plastic F-111 windscreens.

The first question on the observer's rating sheet asked if there was enough distortion in the windscreen to affect flying performance. The second question asked if the windscreen was acceptable (for use on an F-111 aircraft). Results are shown in table 2. Figure 3 plots, against lens factor, the percentage of observers (% I) who say that the distortion is enough to influence flying performance. From the figure it is clear that % I is unrelated to lens factor. The correlation coefficient, r, between LF and % I is only -.1625, too low to be statistically significant. Even when windscreen "M" is excluded, r is only .-1735, also too small to be statistically significant. In either case, the amount of relationship was no more than would be expected by chance.

Figure 4 plots, against displacement grade, the percentage of observers who believe that enough distortion is present to influence flying performance (% I). From the figure it is clear that % I is not related to, or predictable from, DG. This is confirmed by the r between % I and DG of only +.3879 for all 11 windscreens and only -.1737 when windscreen "M" is excluded. Neither value is statistically significant.

Results for the second question on acceptability were similar to the results for influence as far as LF and DG are concerned. The r for LF - % A was only +.1899 for all 11 windscreens and +.1423 for all 11 and +.1899 with "M" excluded. These values are all too low to be statistically significant: no relationship has been shown to exist between DG and acceptability.

Clearly, neither lens factor by itself nor displacement grade by itself was related to either the percentage of observers accepting the windscreens (% A) or to the percentage who expected effects upon flying performance (% I). As mentioned earlier, Thompson had found, in a study of 65 glass F-111 windscreens, that a combination of LF and DG was efficient in sorting windscreens into those acceptable to and those not acceptable to pilots. Earlier in the present paper a quantity R, defined as $R = \sqrt{(LF)^2 + (DG/1{,}000 + 1)^2}$, was shown to do essentially the same sorting as Thompson's acceptability graph. For the 11 windscreens of the present study, the correlation coefficient (r) of % I with R was trivial in magnitude (-.083) and statistically nonsignificant, as Y based on T/N arc sin transform (+.131). Small and statistically nonsignificant r's were also found with windscreen "M" excluded. Clearly, LF and DG in combination, using the "R" formulation, were unrelated to either the acceptability or the judged influence upon flying of the 11 windscreens.

The numbers beneath each data point in figure 2 give, for each windscreen, the percentage of observers who found it acceptable (% A) and the percentage who believed that enough distortion was present to influence

TABLE 2

DISTORTION DATA FOR 11 F-111 WINDSCREENS

	Objective Measures			Overall Subjective Estimates					Subjective Ratings on Specific Types of Distortion					
W*	LF	DG	R+	% I	Arc Sin % I**	% A	Arc Sin % A**	Overall Dist.	Line Split	Line Bend	Line Blur	Banding	Shimmer	Mag.
A	1.40	165	1.821	28.57	32.31	92.86	74.50	1.36	.50	1.21	.71	1.08	.79	.071
B	1.12	120	1.584	28.57	32.31	92.86	74.50	1.36	1.38	1.38	.71	1.43	.57	.64
C	1.10	120	1.570	50.00	45.00	92.86	74.50	1.54	.64	1.57	.57	1.50	.57	.36
D	1.34	140	1.759	14.29	22.21	100.00	90.00	1.29	.57	1.14	.57	1.36	.71	1.43
E	1.04	120	1.528	30.77	33.69	84.62	66.91	1.23	.23	1.00	.31	1.23	.46	.38
F	1.08	120	1.556	21.43	27.58	100.00	90.00	1.29	.43	1.21	.64	1.36	.71	.50
G	1.10	110	1.563	57.14	49.10	78.57	62.42	1.86	.79	1.71	1.00	1.88	.57	.71
H	1.08	100	1.542	7.69	16.10	92.86	74.50	.786	.36	.86	.21	.85	.38	.50
J	1.15	110	1.598	57.14	49.10	64.29	53.27	2.00	.93	1.50	.86	1.93	.62	.57
L	1.04	110	1.521	28.57	32.31	92.86	74.50	1.21	.50	1.29	.57	1.50	.54	.29
M	1.12	170	1.620	92.86	74.50	21.43	27.58	2.64	.86	2.07	1.00	2.57	1.64	1.86
$\bar{X}$	1.143	125.9	1.606	37.64	37.66	83.02	69.33	1.506	.654	1.358	.650	1.515	.687	.665
S.D.	.1178	22.89	.0967	24.41	16.06	22.86	17.44	.4974	.3218	.3418	.2500	.4661	.3366	.5241

*W = arbitrary letter designation of windscreen

** The arc sin transformation to normalize percentage data for data analysis converts percentages to angles: angle = arc sin $\sqrt{(\text{percentage})/100}$

+ R = $\sqrt{(LF)^2 + (DS)/1{,}000 + 1)^2}$ = a composite of lens factor, LF, and displacement grade, DG. LF, DG and R are objective.

I = influence, A = acceptance, Dist. = distortion, Mag. = magnification

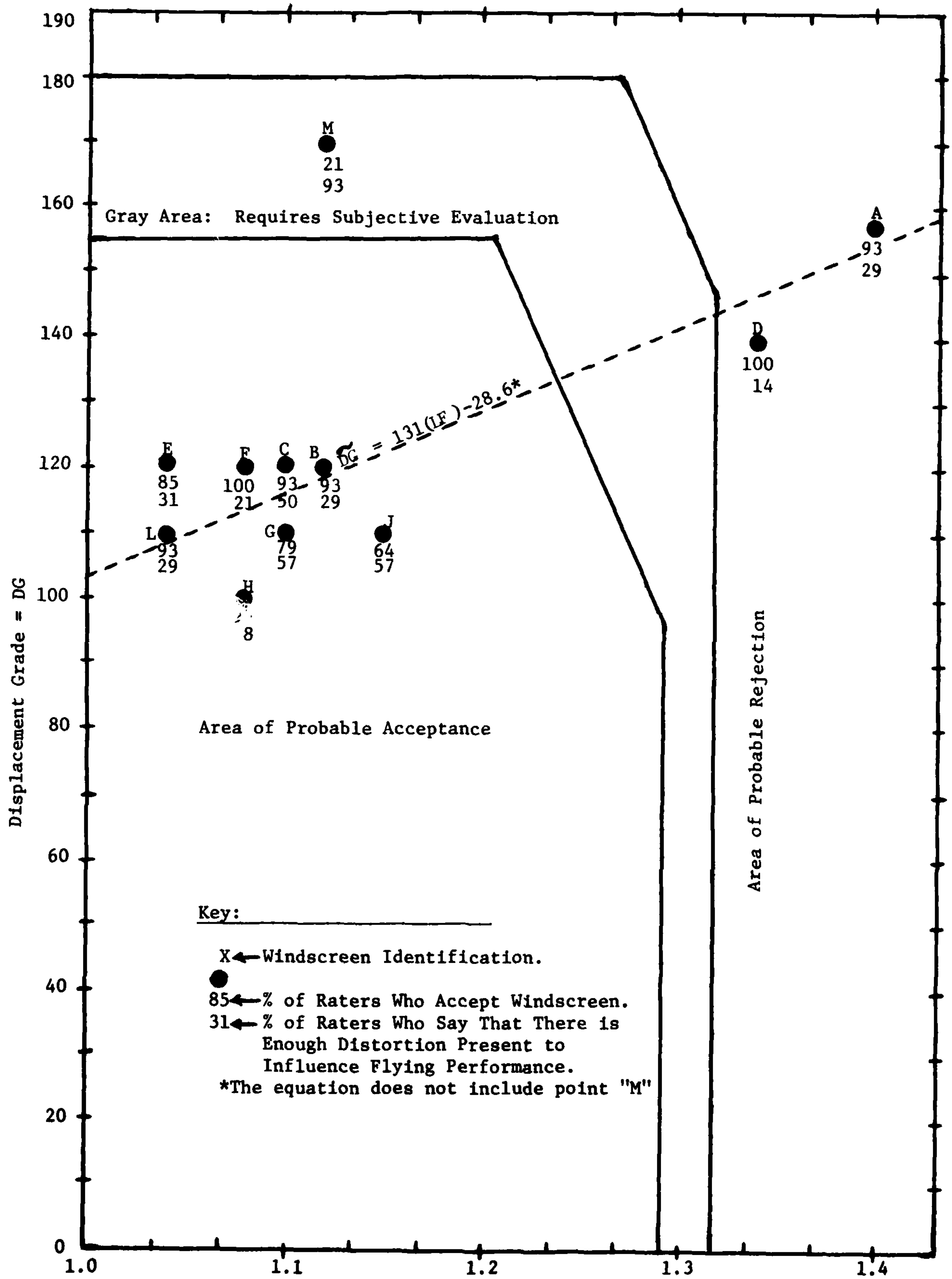


Figure 2. Plot of LF and DG for the 11 F-111 windscreens on a K. E. Thompson Windscreen Acceptability Graph.

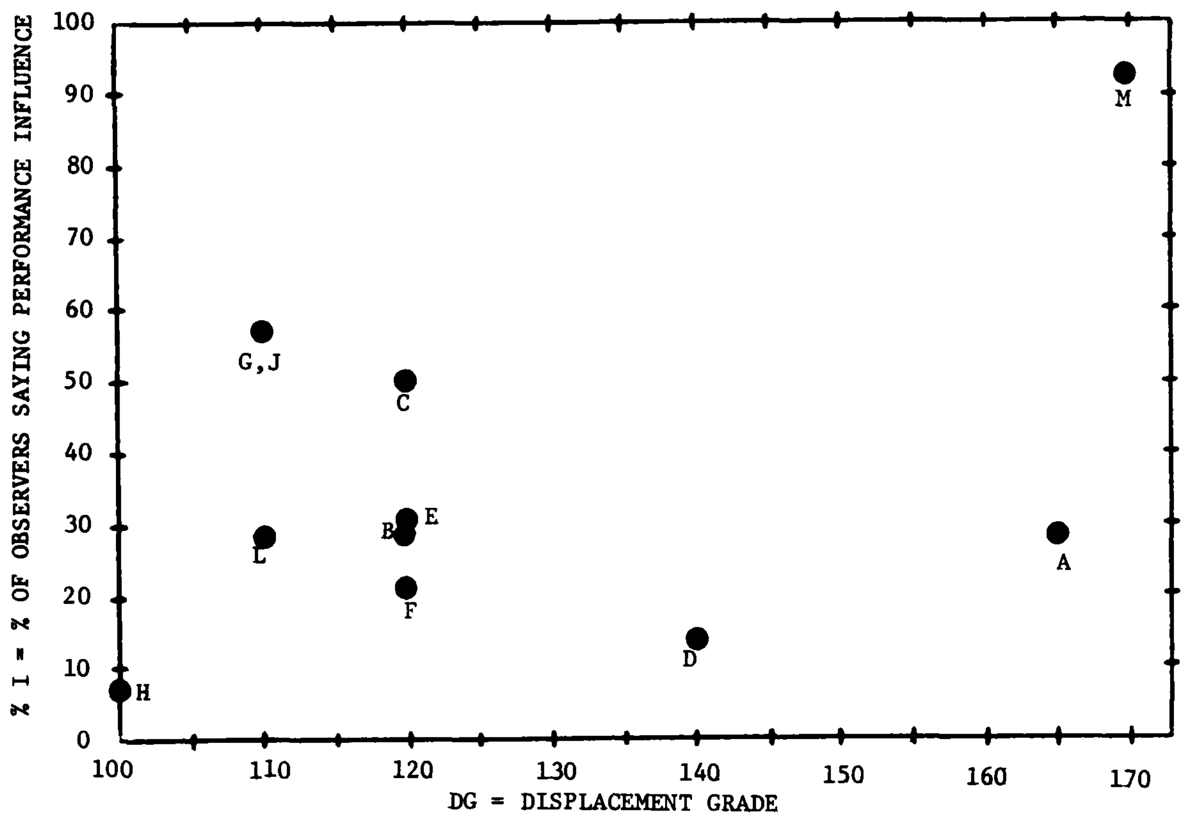


Figure 3. % I plotted against displacement grade.

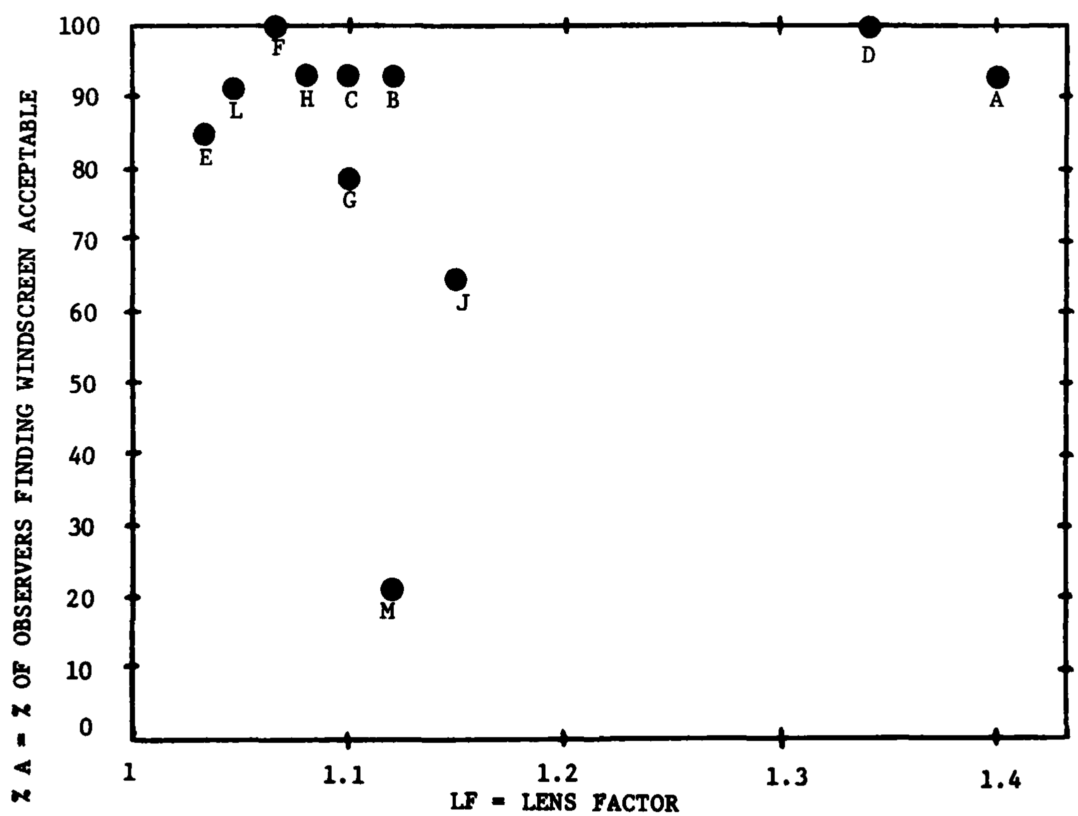


Figure 4. % A plotted against lens factor.

flying performance (% I). This graph's areas of acceptable, questionable (gray) and unacceptable are those of Thompson. Only one of the 11 windscreens was in the gray area (and was unacceptable), while the two in the area of probable rejection were both judged as being acceptable. All windscreens in the area of probable acceptance were found to be acceptable. For the 11 F-111 windscreens that were exclaimed it must be concluded that, like the present author's "R," Thompson's criterion which was also based on LF and DG was not useful for eliminating windscreens that would be optically unacceptable. It is not unlikely that the 11 windscreens that were tested were preselected by the manufacturer to eliminate windscreens that would be unacceptable from the viewpoint of distortion. Even so, doubt is thrown upon the idea that windscreens not passing inspection by Thompson's criteria are unsuitable for installation in aircraft. More data are needed, but it is likely that all windscreens falling into either his gray area or his area of probable rejection will require subjective visual evaluation. Falling into the area of probable rejection, by itself, is not sufficient cause for rejection.

Another inference that can be drawn from figure 2 is that one should not expect to find significant correlations between % A and either LF or DG, nor between % I and either LF and DG. When these four correlations were calculated, none was large enough to attain statistical significance. This is shown in tables 3, 4 and 5.

Displacement guide is correlated positively and significantly with lens factor, as was mentioned earlier, in F-111 windscreens.

ESTIMATED INFLUENCE AND ESTIMATED ACCEPTABILITY

It is to be expected that as the percentage of observers who think that enough distortion is present to influence performance (% I) increases, % A would decrease. The r between influence and acceptability should be high and negative. The correlation tables verify this expectation, with r = -.8916 for all 11 windscreens and r = -.7515 for all except windscreen "M." The r values are not changed much when the data are arch sin transformed, the usual data transformation for percentage data, to make the distributions more nearly Gaussian for statistical analyses.

ESTIMATED INFLUENCE AND DISTORTION ESTIMATES

It is to be expected that % I would be positively correlated with some of the distortion estimates, and the correlation tables bear out this expectation. Statistically significant positive r's are found, when using all 11 windscreens and when omitting the data for windscreen "M," between judged influence on pilot performance (% I) and the overall distortion estimate, grid line bending, grid line blur, banding, and shimmer. Some of the r values are almost .90 and some are even higher. No significant relationship is found between % I and either magnification (or bulls eyes) or line splitting. The range of values of both magnification and of line splitting covered by the 11 F-111 windscreens was quite small, as no windscreen had much of either one. With a restricted range of values, only small values of r between either of these two variables and other variables would be expected.

ACCEPTANCE AND DISTORTION ESTIMATES

As expected, statistically significant correlation coefficients are found, using all 11 windscreens, between % A and the same variables that were significantly correlated with % I. Actual numerical values were somewhat lower. % A, as was the case with % I, did not correlate significantly with estimated magnification nor with grid board line splitting. However, with windscreen "M" omitted, only the estimated amount of banding is significantly correlated with the percentage of observers accepting each windshield (% A).

OVERALL DISTORTION AND DISTORTION ESTIMATES

One would expect that an observer's estimate of the overall distortion in a windscreen would be a mental summation of the various types of distortion that were present. Thus, statistically significant r values with one or more of the various types of distortion would be expected. In the correlation table for all 11 windscreens, there are high positive statistically significant r's between the overall distortion estimate and shimmer, banding, grid line blur and grid line bending. A low but statistically significant r is found between

TABLE 3

PRODUCT MOMENT CORRELATIONS BETWEEN MEASURES FOR ALL ELEVEN F-111 WINDSCREENS

Measure	R+	DG	LF	Mag.	Shim.	Band.	Line Blur.	Line Bend	Line Split	Overall Dist.	%A
Influence (%I)	-.0828	+.3879	-.1625	+.4843	+.7167	+.9292	+.7610	+.9369	+.4170	+.9628	-.8916
%I Arc Sin XFormed	-.0696	+.4211	-.1541	+.4878	+.7388	+.9315	+.7640	+.9379	+.4169	+.9639	-.8836*
Acceptance (%A)	+.0414	-.4510	+.1423	-.6400	+.8091	-.8886	-.5922	-.7624	-.3166	-.8886	
%A Arc Sin XFormed	+.1305	-.3424	+.2089	-.4720	-.6716	-.8006	-.5487	-.7315	-.3340	-.7935	
Overall Distortion	+.0782	+.4799	+.2175	+.6071	+.7893	+.9608	+.8509	+.9317	+.4842		
Line Splitting	+.0308	+.0681	+.0298	+.2985	+.0245	+.5005	+.6598	+.5725			
Line Bending	-.0200	+.3868	-.0907	+.5265	+.7135	+.9300	+.8694				
Line Blurring	+.2168	+.3949	+.1756	+.4131	+.6050	+.8169					
Banding	-.0985	+.3364	-.1700	+.6572	+.7445						
Shimmer	+.3076	+.8070	+.1961	+.7379							
Magnification	+.1763	+.4658	+.1140								
Lens Factor (LF+)	+.9913	+.6311									
Displacement Grade (DG+)	+.7048										

+ R, LF, and DG are objective measures. R is a composite of LF and DG, $R = \sqrt{(LF)^2 + (DG/1000 + 1)^2}$. All other measures are subjective ratings made by observers viewing a grid board through the windscreens.

NOTE: To attain statistical significance at the .05 level the value of r must equal or exceed .521 on a one-tailed test, or .602 on a two-tailed test, since the degrees of freedom are $n - 2 = 11 - 2 = 9$.

* The r between %I and %A when both are arc sin transformed is also -.8836.

The multiple correlation coefficient between estimated overall distortion and shimmer, banding, blurring, bending and magnification is +.8188.

TABLE 4

PRODUCT MOMENT CORRELATIONS BETWEEN MEASURES FOR ALL WINDSCREENS EXCEPT FOR WINDSCREEN "M"

Measure	R+	DG	LF	Mag.	Shim.	Band.	Line Blur.	Line Bend	Line Split	Overall Dist.	%A
Influence (%I)	-.1799	-.1737	-.1735	-.1846	+.0695	+.8392	+.7030	+.8756	+.3971	+.9147	-.7515
%I Arc Sin XFormed	-.1652	-.1299	-.1628	-.2065	+.1097	+.8407	+.7146	+.8788	+.3814	+.9149	-.7085*
Acceptance (%A)	+.2173	+.3464	+.1899	-.1218	+.1921	-.6484	-.4456	-.4476	-.2883	-.7237	
%A Arc Sin XFormed	+.2793	+.3522	+.2606	+.3231	+.3523	-.5098	-.3342	-.4167	-.3220	-.4952	
Overall Distortion	+.0627	-.0069	-.0784	+.0823	+.3526	+.9094	+.8622	+.8652	+.5056		
Line Splitting	+.0208	-.0902	+.0445	+.2154	+.1332	+.5278	+.6156	+.6022			
Line Bending	-.0964	-.0977	-.0645	+.0086	+.2616	+.8613	+.8567				
Line Blurring	+.2192	+.1443	+.2369	+.1068	+.5542	+.8002					
Banding	-.2053	-.2810	-.1849	+.2072	+.1761						
Shimmer	+.7596	+.7822	+.7455	+.1231							
Magnification	+.2177	-.0345	+.2516								
Lens Factor (LF+)	+.9977	+.8752									
Displacement Grade (DG+)	+.9057										

+ R, LF, and DG are objective measures. R is a composite of LF and DG, $R = \sqrt{(LF)^2 + (DG/1000 + 1)^2}$. All other measures are subjective ratings made by observers viewing a grid board through the windscreens.

NOTE: To attain statistical significance at the .05 level the value of r must equal or exceed .549 on a one-tailed test, or .632 on a two-tailed test, since the degrees of freedom are $n - 2 = 10 - 2 = 8$.

* The r between %I and %A when both are arc sin transfored is also -.7085.

The multiple correlation coefficient between estimated overall distortion and shimmer, banding, blurring, bending and magnification is +.4848.

TABLE 5

100X (COEFFICIENT OF DETERMINATION) FOR STATISTICALLY SIGNIFICANT CORRELATION COEFFICIENTS

(A) All eleven F-111 windscreens

Measure	R	DG	LG	Mag.	Shim.	Band	Blur	Bend	Split	Dist*	%A
Influence = %I					51	86	58	88		93	79
%I Arc Sin X'Formed					55	87	58	89		93	78
Acceptance = %A				41	65	74	35	58		79	
%A Arc Sin X'Formed					45	64	30	54		63	
Overall Distortion				37	62	92	72	87			
Line Splitting							44	35			
Line Bending				28	51	86	76				
Line Blurring					37	67					
Banding				43	55						
Shimmer		65									
Magnification											
Lens Factor (LF)	98	40									
Displacement Grade (DG)	50										

(B) All windscreens except windscreen "m"

Measure	R	DG	LF	Mag.	Shim	Band	Blur	Bend	Split	Dist*	%A
Influence = %I						70	49	77		89	56
%I Arc Sin X'Formed						71	51	77		83	50
Acceptance = %A						42				52	
%A Arc Sin X'Formed											
Overall Distortion						83	74	75			
Line Splitting							38	36			
Line Bending						74	73				
Line Blurring				31							
Banding											
Shimmer	58	61	56								
Magnification											
Lens Factor (LF)	99.5	77									
Displacement Grade (DG)	82										

† Coefficients of determination are the squares of the correlation coefficients. When multiplied by 100, they represent the percentage of the total variance in one variable that is associated with or predictable from the variation in the other variable. A value of 100 means that one variable is predictable with zero error from the other variable.

* Dist. = overall distortion estimate.

overall distortion and magnification, but no significant r is found with grid line splitting. According to observer ratings, windscreen "M" contains the most magnification. With the data from "M" excluded, the r between estimated overall distortion and either shimmer or magnification is not statistically significant.

SOME RELATIONSHIPS BETWEEN TYPES OF DISTORTION

The correlation tables give the amount of covariation or correlation between observer ratings, as well as between the objective measures of LF, DG, and their composite, R. Several of these relationships have been discussed, and it was noted that grid board line splitting did not correlate significantly with any other measure, while magnification estimates had correlations that were low, though they were statistically significant in some cases. Overall distortion has already been discussed. This leaves us with the question: What types of distortion have a large impact on % A, % I and overall distortion other than these three themselves? For all 11 F-111 windscreens these distortion types are shimmer, banding, blurring and bending, with shimmer excluded when windscreen "M" is excluded. A question of some interest is: "How closely do these four types of distortion relate to each other?" Table 6 below, for all 11 windscreens, extracts the product moment correlation coefficients from the larger correlation table, table 3.

TABLE 6

CORRELATIONS BETWEEN DISTORTION TYPES FOR ALL ELEVEN F-111 WINDSCREENS

	Shimmer	Banding	Blurring
Bending	+.7135	+.9300	+.8694
Blurring	+.6050	+.8169	
Banding	+.7445		

Note: All tabled r values are statistically significant at the .05 or higher level of significance.

From the table it is clear that: (1) all correlations are positive, as one might expect, (2) all are statistically significant, i.e., the relationship is higher than can reasonably be attributable to chance, (3) all are fairly large in magnitude, with the r's between shimmer and the other three being the lowest. Even the blurring-shimmer r, the lowest, is +.6050, accounting for 36% (or $100r^2$) of variance in prediction.

As indicated in the previous paragraph, the overall distortion rating or estimate by observers is correlated with (or related to) ratings on shimmer, banding, blurring and grid line bending. It may be noted from table 3 for all 11 windscreens that banding ratings correlate +.9608 and bending ratings +.9317 with observer estimates of overall distortion. Thus, either banding or bending ratings can account for most of the variance in overall distortion. In other words, one could use the mean rating on either of these two variables for any given windscreen to do a fair job of predicting the overall distortion rating for the windscreen. The prediction efficiency of shimmer or blur ratings for predicting the overall distortion rating for any of the 11 windscreens is not low, but is lower than for bending or banding estimates: the r for shimmer-overall distortion is +.7893 and the r for blur-overall distortion is +.8509.

The multiple correlation coefficient for overall distortion rating is defined as the correlation between the 11 average values of overall distortion estimates for the 11 windscreens and the best estimate (or prediction) of overall distortion rating for each of the 11 based on the average ratings on several types of distortion. Multiple R is +.9881, a very high and statistically significant value, between overall distortion ratings and predicted overall ratings based on ratings on shimmer, banding, blurring and bending. The square of R is .9763, indicating that 98% of the variance in the overall distortion rating can be accounted for by the ratings on these four variables. The predicted average distortion rating for any one of the 11 windscreens is given by the equation: $Y_i = .2379\ (\text{Shimmer})_i + .6319\ (\text{Banding})_i + .3130\ (\text{Blurring})_i + .1885\ (\text{Bending})_i - .0743$. The values used in the parentheses are the average values for the particular windscreen, i. It is instructive to apply the equation to the 11 windscreens to see how well it predicts for individual windscreens. This is done in table 7.

TABLE 7

OBTAINED AND PREDICTED OVERALL DISTORTION RATINGS

Overall Distortion Rating	WINDSCREEN A	B	C	D	E	F	G	H	J	L	M
Obtained	1.36	1.36	1.54	1.29	1.23	1.29	1.86	.786	2.00	1.21	2.64
Predicted+	1.25	1.45	1.48	1.35	1.10	1.38	1.88	.781	1.84	1.42	2.64
% Error*	+8.1	-6.6	+1.3	-4.7	+10.6	-7.0	-1.1	+6.4	+8.0	-17.4	0

*% Error = (obtained - Predicted) x 100 / (Obtained)

+ The predicted value is based on a combination of shimmer, banding, blurring and bending, using the prediction formula given in the text.

Note from the table that the percentage error in predicting overall distortion rating from a combination of raings on four other variables is small for most windscreens: the two largest errors are 17.4% for windscreen "L" and 10.6% for windscreen "E." The r value between the obtained and predicted values in the table is +.9749 and would be +.9891 had the values not been rounded off to 3 digits. This represents excellent predictivity.

Magnification ratings do not, as indicated earlier, have a high correlation with overall distortion ratings. Even so, a multivariate analysis was conducted with magnification ratings added to those for bending, blurring, banding and shimmer. Multiple R for all 11 windscreens was .9765, accounting for 95.4% of the variance in predicting overall distortion ratings. Multiple R for all windscreens except windscreen M was .9443, accounting for 89.1% of the variance in predicting overall distortion ratings. In both the 10 and 11 windscreen cases the regression coefficient for magnification was small and negative. Adding magnification ratings to the prediction equation did not have much effect upon its prediction efficiency. As a matter of interest, the prediction equations using all five independent variables are as follows for all 11 windscreens and for only 10 ("M" excluded):

Y Overall (11) = .32063 (Shimmer) + .72383 (Banding) + .27942 (Blurring)
+.10848 (Bending) -.09101 (Mag.) -.07977
Y Overall (10) = .33973 (Shimmer) +.72888 (Banding) +.26335 (Blurring)
+.11436 (Bending) -.09055 (Mag.) -.09603

The relative contributions to the overall distortion prediction of the various types of distortion is best seen from examining the coefficients of the various distortion estimates when the equations are in standard form, i.e., relative deviation scores rather than raw scores (or ratings directly from the rating scale) are used. Table 8 gives these coefficients, which are known as standard partial regression coefficients, for the independent (or predictor) variables.

Table 8 shows (1) banding ratings are the most predictive of overall distortion ratings; (2) blurring and shimmer ratings carry considerably less weight than banding, but are comparable to each other; (3) the contribution from bending ratings in predicting overall distortion ratings is quite small; (4) the importance or weight of magnification ratings is not negligible, and exceeds that of bending ratings, and (5) the magnification coefficients are negative, indicating that they have a suppressant effect upon the contribution of the other variables.

TABLE 8

STANDARD PARTIAL REGRESSION COEFFICIENTS FOR PREDICTING OVERALL WINDSCREEN DISTORTION

V*	n**	Shimmer	Banding	Blurring	Bending	Magnification
4	11	.16097	.59204	.15728	.12950	-------
5	11	.48837	.69598	.30266	.10304	-.25547
5	10	.23853	.66522	.18825	.10206	-.25531

V* = number of variables used to predict distortion rating
n** = number of windscreens used

DISTORTION RATINGS OF PILOTS AND QUALITY INSPECTORS

The data discussed up to this point have been for all 14 observers. The observers included six pilots with extensive flying experience and six visual quality inspectors (VQI). The obvious question is: "How closely do the ratings of the two groups compare?" Will windscreens found acceptable by VQI be acceptable to pilots? Ideally, this would be answered by comparing ratings made by VQI's in the factory with ratings made in the field by pilots flying aircraft outfitted with the windscreens in question. This type of data is not available, so what will be compared is ratings made by the two groups in the factory.

The average ratings of the pilot and inspector groups for each of the 11 windscreens is given in table 9. The means or averages for each type of distortion are given at the bottoms of the columns and may be seen in a graphic form in figure 5. Figure 5 shows: (1) Six of the 7 ratings are in very close agreement. (2) The only one not numerically in close agreement, line splitting, is closer to zero distortion than to "Little" distortion for both groups, so that even here, for all practical purposes, agreement is close. (3) Visual quality inspectors rate distortion slightly lower than do pilots on all types of distortion, but the differences between the groups are negligible. From the figure it is clear that the ratings of the two groups have a high positive correlation; they vary together. The product moment correlation coefficient between the two groups on the 7 distortion estimates is +.9557, bearing out this expectation.

TABLE 9

AVERAGE DISTORTION RATINGS FOR PILOTS AND VISUAL QUALITY INSPECTORS

Wind Screen	Line Splitting PILOTS	Line Splitting VQI	Line Bending PILOTS	Line Bending VQI	Line Blurring PILOTS	Line Blurring VQI	Banding PILOTS	Banding VQI	Shimmer PILOTS	Shimmer VQI	Magnification (Bulls Eyes) PILOTS	Magnification (Bulls Eyes) VQI	Overall Distortion PILOTS	Overall Distortion VQI
	Type of Windscreen Distortion Rated by Observers*													
A	.667	.333	1.333	1.333	1.167	.500	.800	1.167	.667	1.167	0	.167	1.167	1.333
B	.667	.167	1.500	1.600	.667	.667	1.333	1.500	.667	.500	1.000	.333	1.333	1.500
C	1.000	.167	1.667	1.667	1.167	.167	1.333	1.667	.833	.500	.500	.333	1.833	1.667
D	.500	.333	1.000	1.167	.333	.667	1.333	1.667	.667	1.000	.167	.167	.667	1.833
E	.167	.200	1.333	.600	.333	.400	1.833	.600	.667	.400	.500	.400	1.833	.800
F	.833	0	1.333	1.333	.833	.500	1.333	1.333	.833	.333	.500	.667	1.500	1.333
G	.833	.167	2.000	1.333	1.000	1.000	2.167	1.500	.833	.500	1.000	.500	2.333	1.667
H	.333	.167	.667	.833	.167	.333	1.000	.800	.500	.333	.667	.500	.833	.833
J	1.286	.500	1.667	1.500	1.000	.500	1.833	2.000	.667	.500	.667	1.167	2.167	1.833
L	.500	.167	1.333	1.167	.500	.500	1.333	1.500	.500	.500	.333	.333	1.333	1.167
M	1.167	.500	2.167	2.000	.667	1.200	2.667	2.500	1.000	2.000	1.833	1.833	2.500	2.833
Sum	7.953	2.701	16.000	14.533	7.834	6.434	16.965	16.234	7.834	7.733	7.167	6.400	17.999	16.799
Mean	.723	.246	1.455	1.321	.712	.585	1.542	1.476	.712	.703	.651	.582	1.636	1.527
Median**	.667	.167	1.333	1.333	.667	.500	1.333	1.500	.667	.500	.500	.400	1.667	1.500
r	+.4757		+.7176		+.0184		+.5690		+.5231		+.7958		+.5124	

*There were 6 pilots and 6 visual quality inspectors, one pilot omitted the Banding rating on Windscreen A, and Shimmer on H, J, and L. One VQI omitted Banding on M. Another VQI had to leave and so did not rate anything on Windscreen E. A third VQI omitted Bending on B.

**The mode is given when multiple equal scores make a proper median uncalculable.

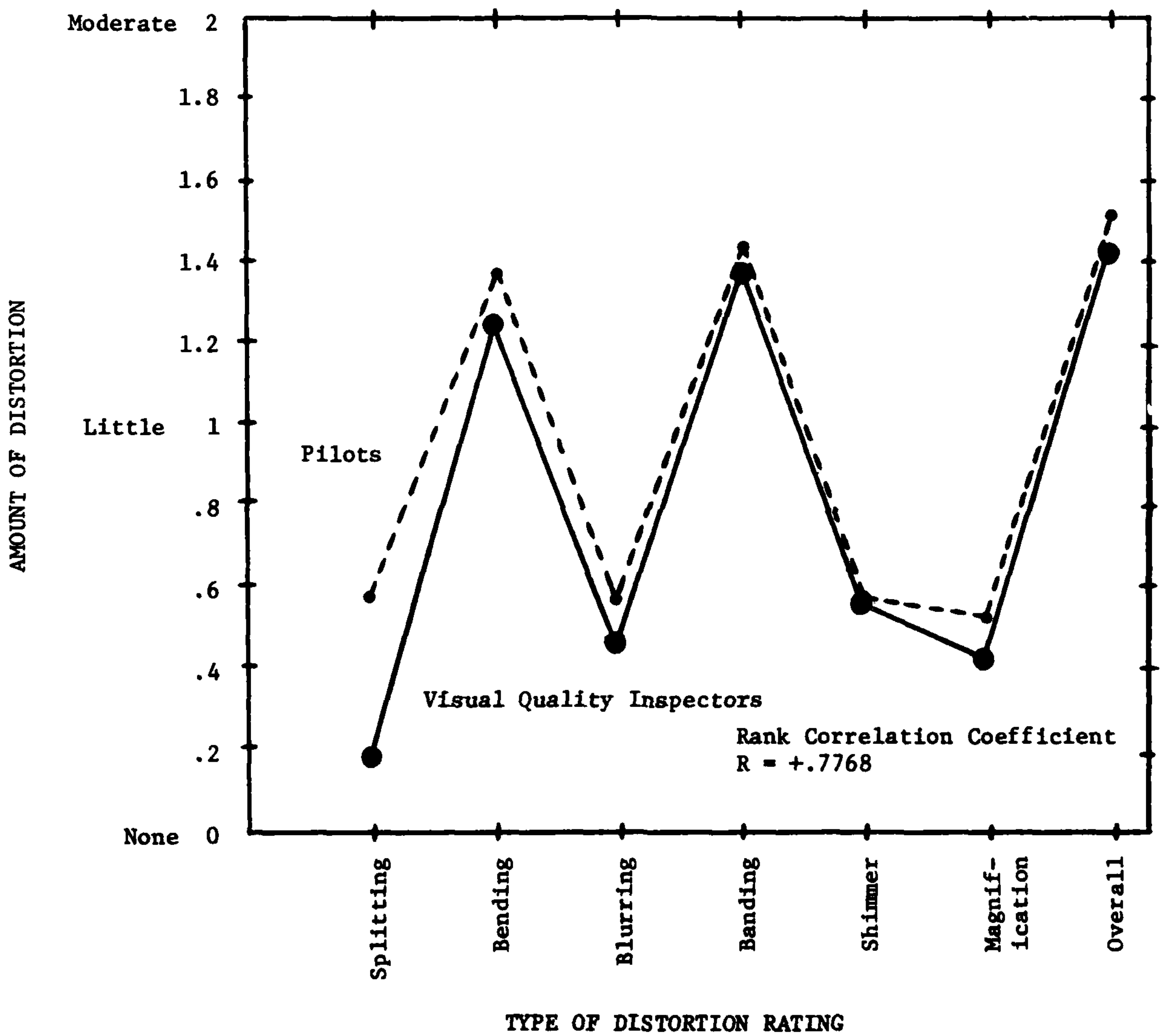


Figure 5. Average Distortion Ratings of 6 Pilots and 6 Visual Quality Inspectors on 11 F-111 Windscreens.

TABLE 10

FREQUENCIES OF USE OF RATING CATEGORIES: (A) BY TYPE OF DISTORTION

Amount of Distortion	Number of Ratings in Each Distortion Category										Distortion Mean*	
	None 0		Little 1		Moderate 2		Large 3		Extreme 4			
Distortion Type	Pilots	VQI	Pilots	VQI	Pilots	VQI	Pilots	VQI	Pilots	VQI	Pilots	VQI
Line Splitting	34	50	18	14	14	1	1	0	0	0	.731	.246
Line Bending	8	7	26	33	25	20	4	4	2	0	1.477	1.328
Line Blurring	27	34	35	22	1	7	2	1	1	0	.712	.609
Banding	7	5	25	27	23	27	9	5	1	0	1.569	1.500
Shimmer	20	32	39	21	4	11	0	1	0	0	.746	.708
Magnification	34	41	25	15	4	5	2	3	1	1	.652	.582
Overall Distortion	3	2	25	34	28	22	9	7	1	0	1.697	1.523
Sum	133	171	193	166	99	93	27	21	6	1	7.584	6.496
Mean	19.00	24.43	25.57	23.71	14.14	13.28	3.86	3.50	.857	.143	1.083	.928
Pilots/VQI	.778		1.163		1.065		1.103		6.000		1.167	
r	+.9799		+.0653		+.7641		+.9089		-----		+.9526	

*Distortion Mean= Σ(number in category)(category mean)/(number in row). Category means are 0, 1, 2, 3, and 4. Example: Line splitting, Pilots' mean=[(34)(0)+(18)(1)+(14)(2)+(1)(3)]/(34+18+14+1)= 49/67=.731

CATEGORIES OF DISTORTION USED BY PILOTS AND INSPECTORS

In the previous paragraph it was shown that pilots and visual quality inspectors were in close agreement on the average amount of each type of windscreen distortion. There were five categories of amount of distortion, ranging from none to extreme. The total number of responses by each group in each distortion category is given in table 10 and is shown in figure 6. It is apparent from the figure that visual quality inspectors tend to use the "None" category more often than the pilots did, with a consequent tendency to use the "Little" category less often. Agreement is close on the "Moderate" and "Large" distortion categories and the magnitude of the difference in the "Extreme" category is small. The frequency of category use may be arranged in the form of a contingency table showing type-of-rater by amount-of-distortion. This is done in table 11.

TABLE 11

NUMBER OF RESPONSES BY DISTORTION CATEGORY FOR PILOTS AND FOR VISUAL QUALITY INSPECTORS

Amount of Distortion	Type of Rater		Sum
	Pilot	VQI	
None	133(.290)*	171(.378)*	304
Little	193(.421)	166(.367)	359
Moderate	99(.216)	93(.206)	192
Large	27(.059)	21(.046)	47
Extreme	6(.013)	1(.002)	7
Sum	458	452	910

*Numbers in parentheses are P values, e.g. 133/458 =.290, etc.

One pilot omitted the banding rating for windscreen A, and the rating for shimmer on H, J and L. One VQI omitted banding on H and blur on M, a second VQI omitted bending on B, and a third VQI was not available to rate E on anything, so the grand sum is 910 instead of 924.

The data in the table yeild a value of chi square of 11.25, with a probability of less than .05. Of the total of 11.25, 8.77 is contributed by the "None" and "Little" categories. It must be concluded that the use frequencies by categories of the two groups are significantly different, with most of the difference being in the two lowest distortion categories, "None" and "Little." Due to the very low usage frequencies, little or no significance should be attributed to the P values for the "Extreme" distortion category. Note that the probability that a pilot will *not* give an extreme rating is .9987(1-.013), and the probability that a VQI will *not* is .9998(1-.002), the difference between the two being trivial.

INDIVIDUAL WINDSCREEN RATINGS BY PILOTS AND VQI

Table 12 gives distortion data for individual windscreens. The means given in the last two columns are plotted in figure 7, with windscreens arranged from left to right in order of increasing distortions as rated by VQI. Note that 9 out of 11 windscreens receiver lower distortion ratings from VQI than from pilots. Figure 7 also shows: (1) the two windscreens with the largest rating difference are actually only a fraction of a distortion category apart, and (2) no trend is obvious for the differences in ratings to be related to amount of distortion as perceptually judged. Thus, agreement is not closer for low distortion windscreens than for high distortion ones, or the reverse, etc.

A more sensitive examination for a differential trend may be made by estimating the distortion of each windscreen as the average of the two groups, and plotting against this the distortion ratings of the two groups. This is done because no satisfactory criterion of absolute amount of distortion is available. Even though such a procedure verges upon "bootstrapping," the results can be illuminating. Figure 8 plots the distortion ratings of pilots and VQI against the group means used as a measure of amount of distortion. The least-squares best fitting straight lines are shown. If both groups were the same, the equations would both be Y = X. Note that neither is very far from this. The lines are close together and do not depart much from being

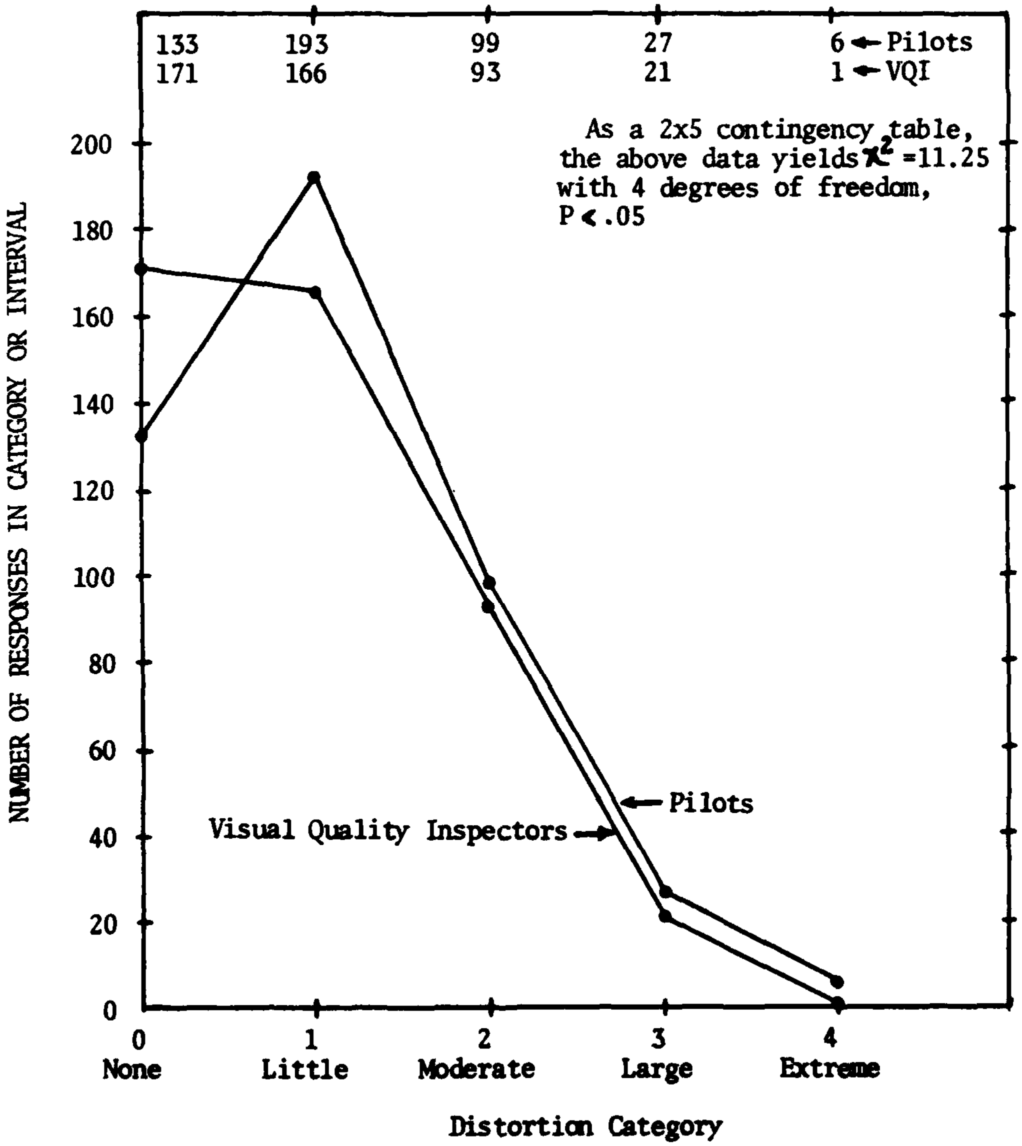


Figure 6. Reponse frequencies in various intervals of the distortion scale for pilots and visual quality inspectors.

TABLE 12

FREQUENCIES OF USE OF RATING CATEGORIES: (B) BY INDIVIDUAL WINDSCREEN+

Amount of Distortion	Number of Ratings in Each Distortion Category												Distortion Mean*	
	None 0		Little 1		Moderate 2		Large 3		Extreme 4		Row Sum			
Windscreen	Pilots	VQI	Pilots	VQI	Pilots	VQI	Pilots	VQI	Pilots	VQI	Pilots	VQI	Pilots	VQI
A	19	14	9	20	11	8	1	0	1	0	41	42	.927	.857
B	9	16	23	17	10	5	0	3	0	0	42	41	1.024	.878
C	7	17	22	13	11	12	2	0	0	0	42	42	1.190	.881
D	18	17	16	11	8	12	0	2	0	0	42	42	.762	.976
E	15	18	16	17	9	0	2	0	0	0	42	35	.952	.486
F	9	17	24	18	8	7	1	0	0	0	42	42	1.024	.762
G	6	14	19	17	10	10	6	1	1	0	42	42	1.452	.952
H	19	21	20	18	1	2	1	0	0	0	41	41	.610	.537
J	9	13	16	10	7	18	9	0	0	1	41	42	1.390	1.190
L	16	16	15	20	10	6	0	0	0	0	41	42	.854	.762
M	6	8	13	5	14	13	5	15	4	0	42	41	1.714	1.854
Sum	133	171	193	166	99	93	27	21	6	1	458	452	11.899	10.135
Mean	12.09	15.55	17.54	15.09	9.00	8.45	2.45	1.91	.545	.091	41.64	41.09	1.082	.921
P/VQI	.778		1.163		1.065		1.286		6.00		1.013		1.174	
Chi Square**	9.36		14.05		18.09		-----		----		----		+.8002	

*Distortion Mean=$\Sigma(fX)/\Sigma f$, where f=frequency and X=category mean.

+Numbers are for splitting, bending, blurring, banding, shimmer, magnification (or bulls eyes), and overall distortion.

**For Chi Square to attain statistical significance at the .05 level with 10 degrees of freedom, its value must equal or exceed 18.31. None did so. Too many zero and low numbers in the large and extreme categories invalidate chi square for them, so it is not given for those categories.

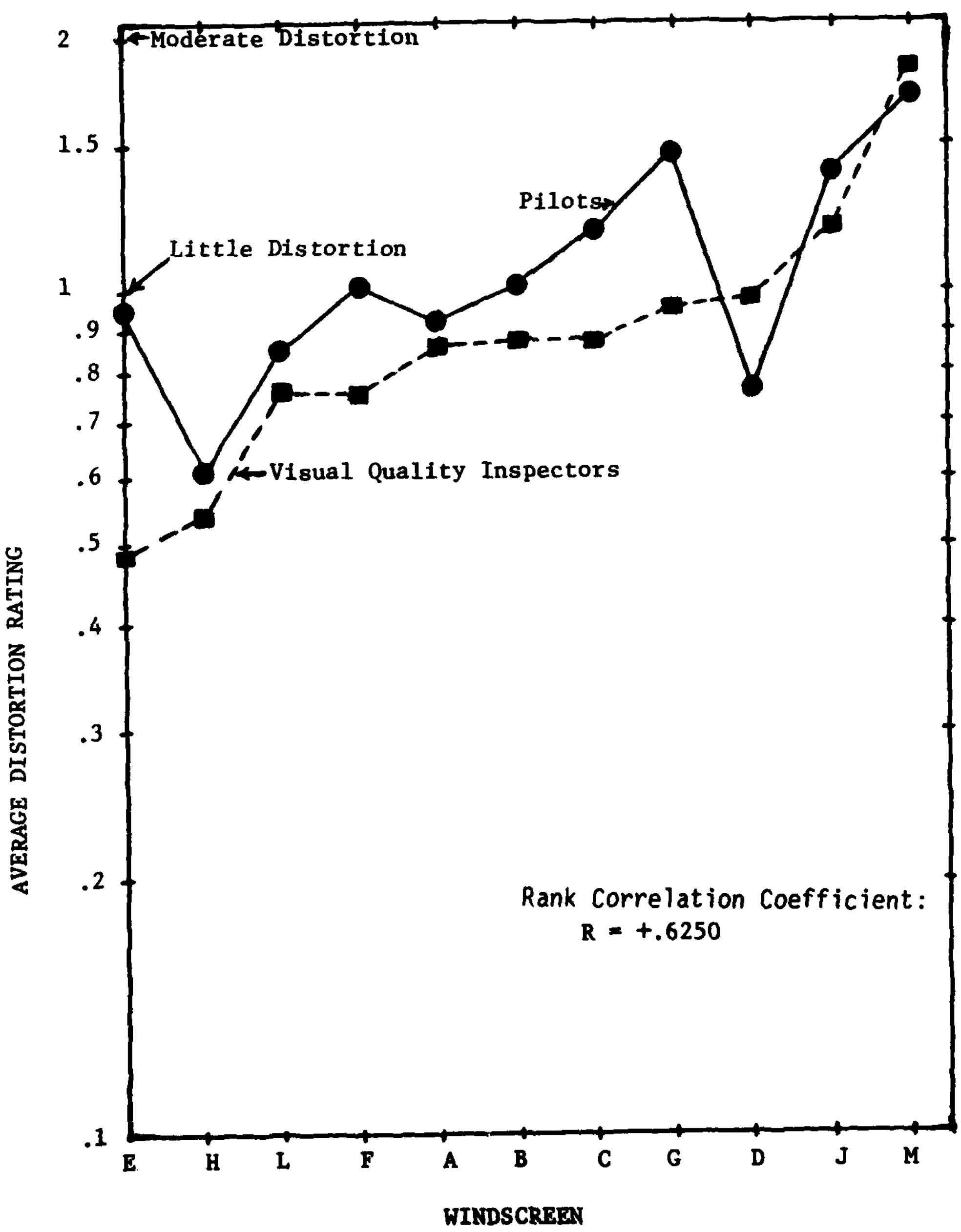


Figure 7. Average Distortion Ratings of 11 F-111 Windscreens for Pilots and for Visual Quality Inspectors. Windscreens are Arranged in Order of Increasing Distortion as Rated by VQI.

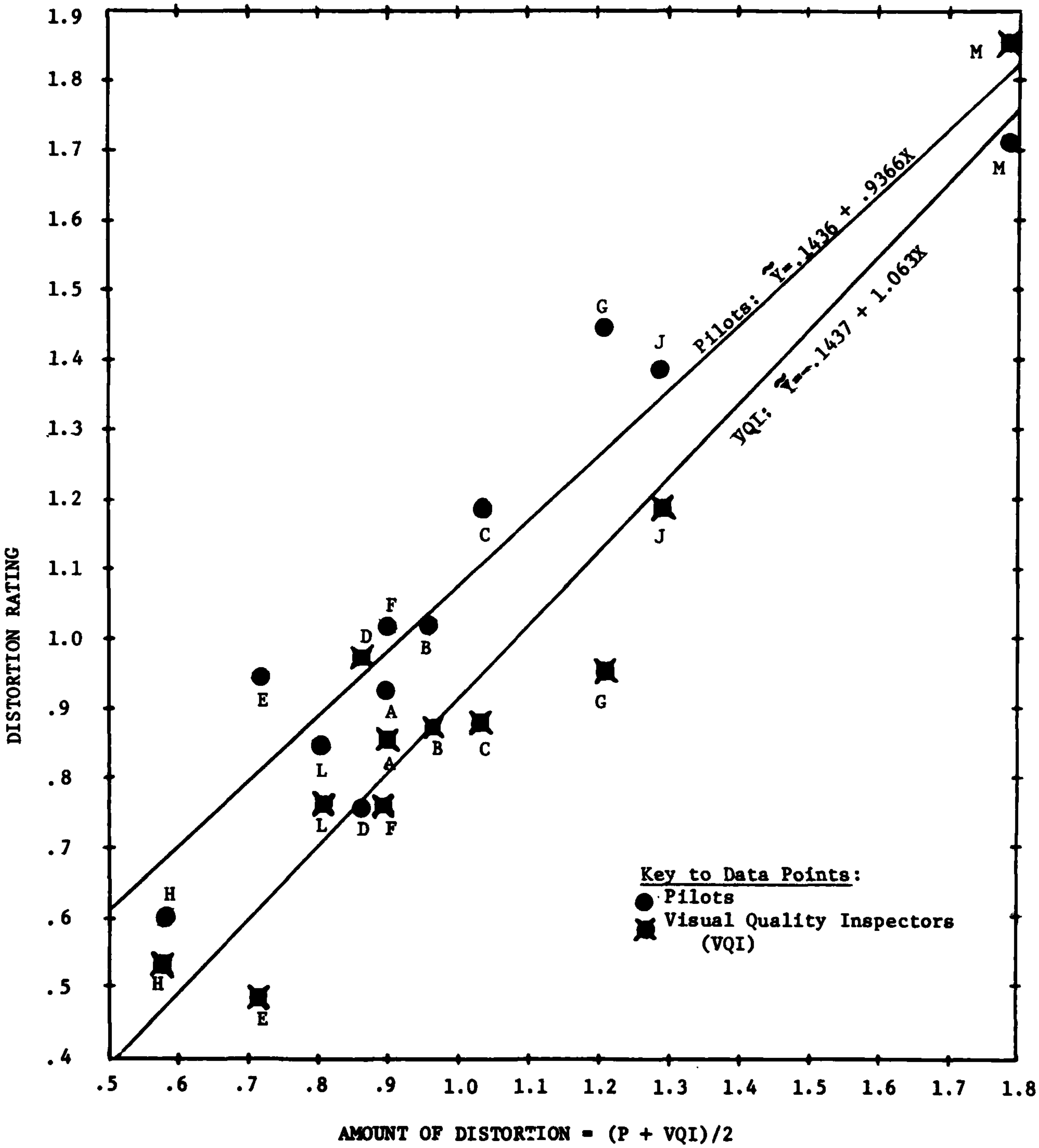


Figure 8. Distortion ratings for pilots and inspectors plotted against amount of distortion estimated as the pilot-inspector average.

parallel. The slight tendency for the curves to converge (to indicate closer agreement) at higher levels of distortion is too small to be reliable. Over the restricted range of distortion afforded by the 11 F-111 windscreens, no reliable evidence indicates any change in amount of agreement with amount of distortion as perceptually judged. In particular, VQI do not become more lenient or less critical as compared to pilots at higher distortion levels. They are only slightly less critical than pilots at all levels that were examined in the present study.

PILOTS AND INSPECTORS ON INFLUENCE AND ACCEPTABILITY

The numbers of pilots and inspectors who believe that each windscreen will or will not influence pilot performance and is or is not acceptable are given in table 13. The bottom line of the table gives percentage values. Note that 59% of the pilots said that flying performance would be affected while 77% said yes on acceptability. This may be contrasted with the 15% of visual quality inspectors on influence and 92% on acceptability. It is clear that pilots were more likely than VQI to believe that the distortion that was present would influence flying performance and were also more likely than VQI to give an unacceptable rating to a windscreen. The VQI were less critical than pilots in judging both influence and acceptability.

Tables 14 and 15, set up as contingency tables, contain the column totals for the frequencies in table 13. The values of chi square calculated from the numbers in these tables were 24.88 ($P<.01$) for influence on flying performance and 4.79 ($P<.05$) for windscreen acceptability. The finding that inspectors were less critical than pilots in judging both the effects of distortion on flying performance and windscreen acceptability is thus supported by statistical significance tests.

TABLE 13

PERFORMANCE INFLUENCE AND ACCEPTABILITY

Screen	Will This Windscreen Affect Flying Performance?						Is This Windscreen Acceptable?					
	Pilots			VQI			Pilots			VQI		
	Yes	No	% Yes	Yes	No	% Yes	Yes	No	% Yes	Yes	No	% Yes
A	4	2	66.7	0	6	.0	5	1	83.3	6	0	100.0
B	2	4	33.3	1	5	16.6	5	1	83.3	6	0	100.0
C	5	1	83.3	2	4	33.3	5	1	83.3	6	0	100.0
D	1	5	16.7	0	6	.0	6	0	100.0	6	0	100.0
E	3	3	50.0	0	5*	.0	4	2	67.7	5**	0	100.0
F	3	3	50.0	0	6	.0	4**	1	80.0	6	0	100.0
G	6	0	100.0	1	5	16.7	4	2	66.7	6	0	100.0
H	1	5	16.7	0	6	.0	5	1	83.3	6	0	100.0
J	5	1	83.3	1	5	16.7	5	1	83.3	5	1	83.3
L	3	3	50.0	0	6	.0	5	1	83.3	6	0	100.0
M	6	0	100.0	5	1	83.3	2	4	33.3	2	4	33.3
Sum	39	27	650.0	10	55	166.6	50	15	847.5	60	5	1016.6
%	59.09	40.91	59.1	15.38	84.62	15.1	76.92	23.08	77.0	92.31	7.69	92.4

*n=only 5, not 6, for VQI on windscreen "E"
**n=only 5, not 6, for VQI on windscreen "E" and for pilots on windscreen "F"
Variation in n causesslight differences in the percentage values at the bottom of the Yes and % Yes columns.

TABLE 14

PERFORMANCE INFLUENCE

Rater	Affect Performance		Sum
	Yes	No	
Pilots	39	27	66
VQI	10	55	65
Sum	49	82	131

Chi Square=24.88, P<.01
Significantly more pilots than VQI reject windscreens.

TABLE 15

ACCEPTABILTY

Rater	Acceptable		Sum
	Yes	No	
Pilots	50	15	65
VQI	60	5	65
Sum	110	20	130

Chi Square=4.79, P<.05
Significantly more pilots than VQI believe performance will be affected.

SECTION IV

DISCUSSION OF RESULTS

In the present study the observers were not trained to make distinctions between types of distortion nor were types of distortion defined or described. It is likely that clear-cut distinctions were not always drawn by observers between types of distortion. Similarly, no training was given on the meaning of terms such as "Moderate Distortion." Some observers may have tried to make absolute judgements while others may have been influenced by the distortion levels of the particular set of 11 windscreens. They were already familiar with them from participating in a separate study.

The above lack of refinement or sophistication is due in part to the lack of a set of standards that would clearly distinguish both types and amounts of windscreen optical distortion. For types of distortion that do not require motion to detect, the standards could be a set of photographs made by photographing a grid board through a distorting panel or computer processing an undistorted photograph to obtain precisely controlled types and amounts of distortion. For a dynamic type of distortion, such as shimmer, a set of three-dimensional holograms would probably be required.

In the present study there were high correlations between types of distortion. To what extent are these relationships or correlations generalizable? This question cannot be answered because the correlations could have been the result of such things as (1) confusion in observers on distinctions between types of distortion, (2) peculiarities of the set of 11 F-111 windscreens due to materials or manufacturing processes. It is hypothesized that even extensive observer training with a set of excellent optical distortion standards would have little, if any, effect upon reducing the high correlations between types of distortion. Types of distortion readily definable and commonly regarded in theory to be distinct or separate may not be independent in realizable windscreens.

Aircraft pilots and visual quality inspectors were quite close in their estimates of both overall optical distortion and specific types of distortion. While pilots rated distortions slightly higher than did the inspectors, the differences were inconsequential. The equivalence of the ratings, particularly at higher levels of distortion where differences could have important practical consequences, is encouraging. It is probably safe to conclude that, in general, subjective estimates of amount of distortion by visual quality inspectors and by aircraft pilots will not differ by much.

The equivalence in estimations of amount of distortion is not matched by equivalence in judging whether or not enough distortion is present to influence flying performance. Pilots were appreciably more likely to rate a windshield as likely to influence performance. Actually, neither the C-141 pilots nor the visual quality inspectors have had any experience with, or even seen data on, situations in which flying performance was judged or measured when optical distortion levels were varied. Thus, it cannot be said that the pilots were too severe in their judgment of probable influence on flying or that the inspectors were too lenient.

Windscreens that do not significantly degrade pilot performance are required. However, it is not known just how high optical quality might be before additional quality is not accompanied by signficant increases in the performance of aircraft pilots. Very high optical quality is both extremely expensive and unnecessary. A level of optical quality can be found that is entirely adequate and is financially affordable. What the Air Force needs now are valid specifications of windscreen optical characteristics and inspection methods that will insure that specifications are met. The present paper has identified several optical characteristics that appear to pilots and windscreen inspectors to be relevant to pilot performance. Further research on pilot requirements and on objective and subjective methods for specifying and measuring windscreen optical qualities are needed.

★U.S. GOVERNMENT PRINTING OFFICE: 1982—559-006/2014

ATE

LME

8